Two crystal structures of N-acetyltransferases reveal a new fold for CoA-dependent enzymes

Yorgo Modis[1] and Rik Wierenga[2]*

Coenzyme A (CoA) is well known for its importance in various metabolic pathways. The recent determination of the structures of several CoA-dependent N-acylating enzymes highlights the importance of the acyl-CoA molecule for intracellular regulation and reveals a novel fold. The N-acetyltransferase fold defines yet another protein superfamily and adds to the already diverse world of CoA-dependent enzymes.

Addresses: [1]European Molecular Biology Laboratory, Postfach 10.2209, D-69012, Heidelberg, Germany and [2]Department of Biochemistry, University of Oulu, Oulu, FIN-90570, Finland.

*Corresponding author.
E-mail: rik.wierenga@oulu.fi

Structure 15 November 1998, 6:1345–1350
http://biomednet.com/elecref/0969212600601345

Introduction

The cofactor coenzyme A (CoA; Figure 1) is the predominant acyl-group carrier cofactor in living systems [1]. Acyl groups are linked to CoA by a thioester bond between the SH group of CoA and the carboxyl group of the acyl moiety. The high energy of the thioester bond makes thioesters less stable and more reactive than oxygen esters. This reactivity facilitates both nucleophilic reactions at the carbonyl center and electrophilic reactions at the adjacent α-position [2]. There is a wide range of possible reactions involving nucleophilic attack at the carbonyl center resulting, for example, in the formation of an acid (reaction with a water molecule, catalysed by an esterase), a peptide bond (reaction with a free amine, catalysed by an amine acyltransferase) or an acetoacyl molecule (reaction with an acetyl-CoA molecule, catalysed by a thiolase). In these three examples the C–S bond is replaced by a C–O, C–N or C–C bond, respectively. The higher reactivity at the α-carbon is exploited by several enzymes, including enoyl-CoA hydratase and citrate synthase. The reactions catalysed by these two enzymes are facilitated by the relatively high acidity of the α proton.

Acyl-CoA molecules are used in several degradative pathways. For example, the degradation of fatty acids in the β-oxidation pathway involves a sequence of four reactions in which the substrates are fatty acyl-CoA molecules. The most prominent acyl-CoA molecule is acetyl-CoA. Acetyl-CoA provides the starting point for various synthetic pathways, such as the synthesis of polyketides [3], and is also the end product of several degradative pathways. For example, the degradation of fatty acids (through the β-oxidation pathway) and carbohydrates (through glycolysis) results in the formation of acetyl-CoA, which is subsequently metabolised further via the Krebs cycle. In addition to its function as a metabolic intermediate, it has been established that acetyl-CoA is also involved in important regulatory modifications in eukaryotic transcription processes [4].

What structural information is available for enzymes that catalyse acyl-CoA-dependent reactions? The first structure of a CoA-binding enzyme, citrate synthase, was solved in 1987. Since then, the structures of 19 more CoA-binding proteins have become available (Table 1). Most of these proteins are enzymes are involved in cellular metabolism, but the recently determined structures of yeast N-myristoyl transferase (NMT) [5] and yeast histone acetyltransferase 1 (Hat1) [6] point to important regulatory roles that can be attributed to acyl groups. The reaction catalysed by NMT modifies the N-terminal glycine residue of several proteins; this modification is essential for the growth of yeast. Hat1 catalyses the acetylation of specific lysine sidechains of histones. This modification is essential for the proper assembly of histones and DNA into nucleosomes, and the enzyme is therefore important in the regulation of transcription processes. Both these enzymes are CoA-dependent and both transfer an acyl group to a free amine group. Hat1 belongs to a large family of enzymes, referred to as the GCN5-related N-acetyltransferases (the GNAT family).

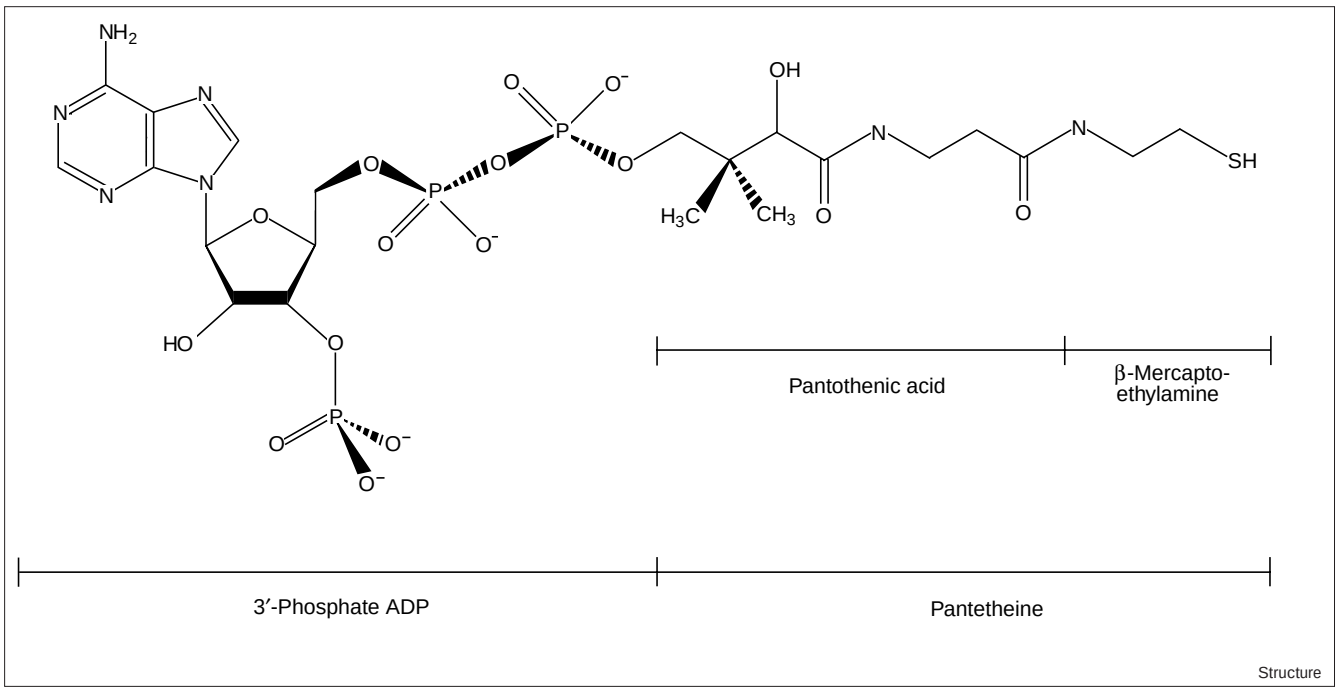
The GNAT family of enzymes

This group of acetyl-CoA-dependent N-acetyltransferases consists of more than 150 proteins with conserved sequence motifs [7]. Enzymes from the GNAT family have diverse functions, which include histone acetylation, transcription factor activation, amino acid metabolism, drug resistance and detoxification. Until recently, no structural data was available for this class of enzymes. In the August 1998 issue of *Cell*, however, the structures of two enzymes belonging to different subclasses of this family were described. These articles reported the crystal structures of the GCN5-related aminoglycoside 3-N-acetyltransferase (AAT) [8] and Hat1 from yeast [6].

Histone acetyltransferase 1 (Hat1)

The crystal structure of yeast (*Saccharomyces cerevisiae*) Hat1 (residues 1–320; the C-terminal 54 residues were deleted) in complex with acetyl-CoA has been solved at 2.3 Å resolution [6]. The core of the structure comprises

Figure 1



The covalent structure of coenzyme A. The various building blocks that make up CoA are shown. The ADP moiety also occurs in other nucleotide cofactors, such as FAD and NAD, which are used by many enzymes involved in oxidation/reduction reactions. The pantetheine moiety by itself is also used by other enzymes, for example those involved in fatty acid synthesis [1] and nonribosomal protein synthesis [15].

residues 140–270. Residues from the N terminus to residue 140 form an α-helical domain; the remaining residues at the C terminus complete the core domain. The secondary structure of the core domain is shown in Figure 2, together with the mode of binding of acetyl-CoA. The core domain consists of a β sheet covered on both sides by α helices. The sheet is made up of a mixture of parallel and antiparallel β strands. Most important for the interactions with acetyl-CoA is the βαβ unit at the end of the sheet. The pyrophosphate moiety of CoA interacts with the loop leading into the α helix of the βαβ unit (the loop after β strand S4 in Figure 2a). This mode of binding is reminiscent of the classical mode of binding of dinucleotides to the Rossmann fold [9], as shown in Figure 2b. It should be noted, however, that the directionality of the nucleotide is reversed: the relative position of the adenine moiety which is 'above' the sheet in Hat1 is 'below' the sheet in the classical mode of binding. So far, only one CoA-dependent enzyme has been shown to have a CoA-binding domain with a Rossmann fold, namely succinyl-CoA synthetase [10]. The mode of CoA binding of this enzyme is shown in Figure 2c. An important feature of the Rossmann fold is the tight turn between the first β strand and the α helix of the βαβ unit, which is characteristic of NAD- and FAD-binding Rossmann folds. This loop has a characteristic sequence motif, Gly-x-Gly-x-x-Gly [11]. In Hat1 and other GNAT family members, this loop is rather long. Interestingly, this loop also has a very conserved sequence signature in the GNAT family of enzymes, referred to as motif A (Arg/Gln-x-x-Gly-x-Gly/Ala) [7]. The loop with sequence motif A binds the pyrophosphate group, the central moiety of the acetyl-CoA molecule. In the Hat1 structure this part of the acetyl-CoA molecule is well defined in the electron-density map. The adenosine moiety of acetyl-CoA is less well defined in the structure, but the acetyl-pantetheine moiety, extending downwards with respect to the pyrophosphate moiety in Figure 2a, is well defined. The methyl group of the activated acetyl group is buried at the bottom of a hydrophobic pocket. This hydrophobic pocket is formed by residues near the C-terminal ends of the two parallel β strands of the βαβ unit (e.g. Ile217 and Val254).

In addition to motif A, three other sequence motifs have been described to be characteristic for the GNAT family of enzymes: motifs B, C and D. Motifs C and D seem to be more related to the fold, in that residues of these motifs contribute to the hydrophobic core. Motif B is close to the active-site pocket and includes residues following the second parallel β strand (β strand S5; in Figure 2a).

Table 1

Published structures of CoA-binding enzymes.

Enzyme	PDB entry code	Reference
Citrate synthase	2cts	[17]
Dihydrolipoyl transacetylase	1eab	[18]
Acyl-CoA dehydrogenase	3mde	[19]
Succinyl-CoA synthetase	1scu	[10]
Methymalonyl-CoA mutase	1req	[20]
Enoyl-CoA hydratase	1dub	[21]
Acyl-CoA-binding protein	1aca	[22]
3-Hydroxyacyl-CoA dehydrogenase	–	[23]
Choramphenicol acetyltransferase	1cla	[24]
3-Ketoacyl-CoA thiolase	1afw	[13]
3-HMG-CoA dehydrogenase	–	[25]
Malonyl-CoA ACP transacylase	1mla	[26]
4-Chlorobenzoyl-CoA dehalogenase	1nzy	[27]
Isovaleryl-CoA dehydrogenase	1ivh	[28]
Lipid-transfer protein	1jtb	[29]
Dienoyl-CoA isomerase	1dci	[30]
N-myristoyl-CoA transferase	1nmt	[5]
Xenobiotic acetyltransferase	2xat	[31]
Aminoglycoside 3-N-acetyltransferase	–	[8]
Histone acetyltransferase Hat1	1bob	[6]

This table includes structures described in the literature before October 1998. Only one Protein Data Bank (PDB) entry code is given for each enzyme, even when more are available.

Hat1 is a monomeric enzyme; however, its *in vivo* activity is regulated by complex formation with another protein called Hat2. Thus, the extra residues before and after the acetyl-CoA binding core domain may be important for interacting with Hat2.

GCN5-related aminoglycoside 3-N-acetyltransferase (AAT)

The structure of bacterial AAT (residues 1–168, excluding residues 169–177) in complex with CoA, has been solved at 2.3 Å resolution [8]. AAT catalyses the N-acetylation of aminoglycoside antibiotics. This modification causes the bacterium to become resistant to these antibiotics, which explains the considerable clinical interest in this class of enzymes. *In vivo*, AAT is most likely to be a dimer of two identical subunits of 177 residues each. In the crystal structure the asymmetric unit contains one dimer, with tight interactions between the two subunits. Each subunit is equivalent to the core of the CoA-binding domain of Hat1 and has the same topology as the Hat1 core domain. The sequence identity between these two domains is very low: a sequence alignment of the two domains results in a sequence identity of 9%. The mode of CoA binding observed in the AAT complex is the same as that seen in the Hat1 complex (Figure 2a). The CoA molecule is not bent at the pyrophosphate moiety, as seen in some other complexes [12]. Instead, the pantetheine moiety adopts a bent conformation, such that the methyl groups of the pantothenic acid point to the β-mercaptoethylamine moiety.

N-myristoyl transferase (NMT)

The structure of unliganded yeast (*Candida albicans*), N-myristoyl-CoA transferase (NMT) has also been solved recently, at 2.8 Å resolution [5]. This monomeric enzyme of 451 residues catalyses the transfer of the fatty acid myristate (C14:0) from myristoyl-CoA to the N-terminal glycine residue of a number of proteins. In yeast this is an essential enzyme and is therefore an attractive target for potential antifungal pharmaceuticals.

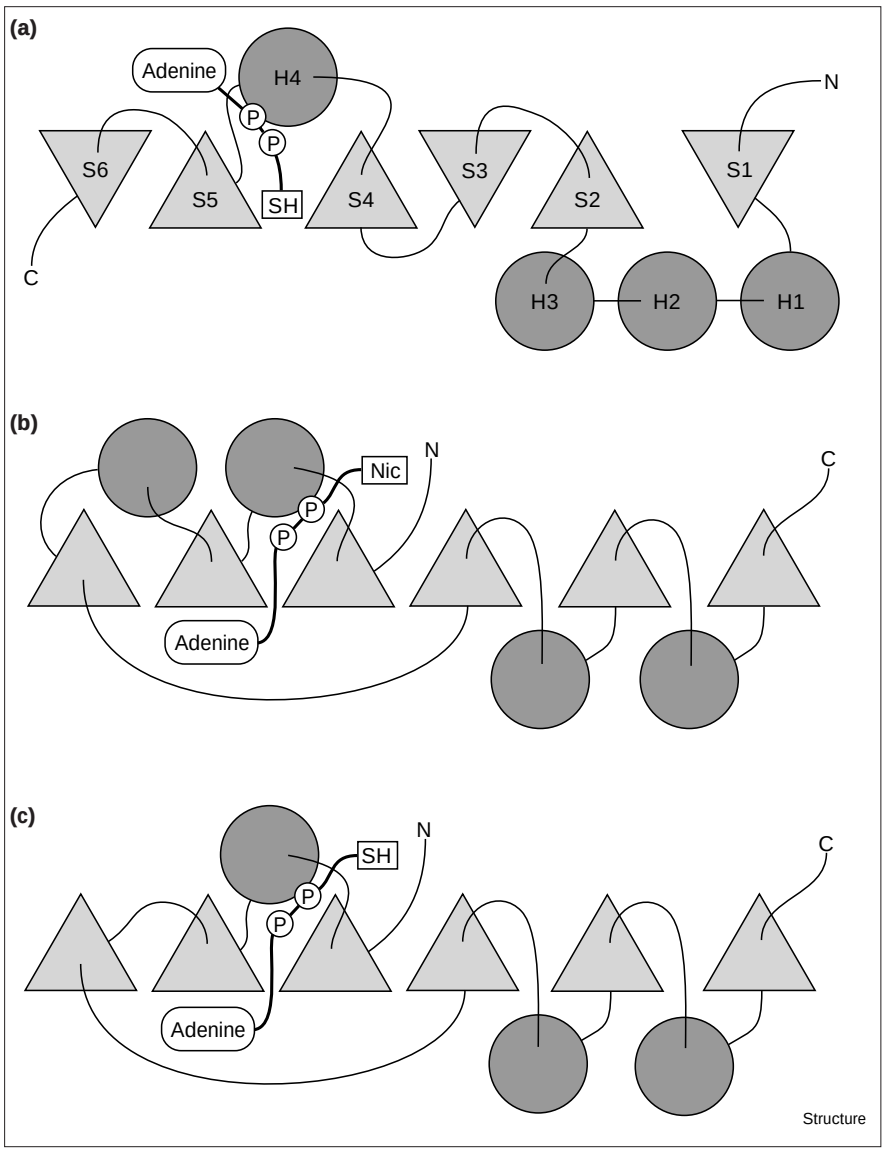
As only the structure of the unliganded enzyme has been described, the location of the active site can only be deduced from the surface features of the structure and from sequence conservation and mutational studies. The crystal structure shows the molecule to be made up of two halves with the same topology; this topology is essentially the same as that of the core domain of the GNAT family of enzymes discussed above (Figure 2a). It remains to be seen if this topological similarity extends to a functional similarity, but it is interesting to note that the NMT family of enzymes also contain a conserved sequence motif in the loop between β strand S4 and α helix H4 (see Figure 2a). This loop is positioned near the presumed active site of NMT.

Aspects of the reaction mechanism

It seems that each of the three enzymes discussed above binds to acyl-CoA first, and subsequently binds to the substrate. In each enzyme the acyl group is transferred directly to the free amine and not via a protein SH group as is the case in thiolase, for example [13].

In Hat1 and AAT, the active site is located near the C-terminal end of the parallel β strands 4 and 5 (Figure 2a). Some residues of these β strands contribute to the active-site pocket. This is possible because the two strands do not form a tight βαβ fold, as seen in the Rossmann fold, but instead run apart from each other leading to the formation of an active-site pocket (Figure 3). At present, it is not possible to draw detailed conclusions about the reaction mechanism as none of the current structures provide information about the ternary complex. Each of these enzymes must have a mechanism to deprotonate the free amine. It is speculated that the acidic residues of motif B in AAT are important for neutralising the protonated free amine group [8]. Similarly, it has been suggested that conserved

Figure 2



The topology of the CoA-binding region of members of the GNAT family compared to the classical Rossmann fold nucleotide-binding domain. **(a)** The secondary structure elements of AAT (residues 26–161); β strands are represented as triangles and α helices as circles. The nomenclature is as used originally [8]. In the CoA-binding domain of Hat1 (residues 140–270) the last strand of the β sheet is replaced by an α helix. **(b)** The mode of binding of NAD to the classical Rossmann fold, as seen in lactate dehydrogenase [9]. **(c)** The mode of binding of CoA in succinyl-CoA synthetase [10].

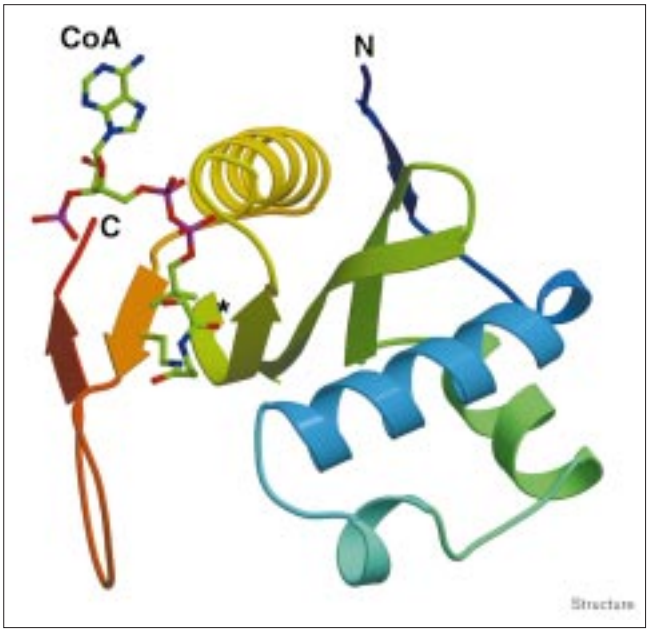
acidic residues near the active site (e.g. Glu173 in β strand S4) might have this function in NMT [5].

In the acetyltransferase reaction, the activated acyl group is transferred to a free amine; the product is therefore a peptide bond. It is interesting to note that a similar reaction is used in nonribosomal protein synthesis. Peptide bond formation in nonribosomal protein synthesis is carried out by large multienzyme systems [14] that catalyse the formation of a variety of bioactive peptides. These multienzyme systems consist of many domains, one of which catalyses peptide bond formation. For this reaction, the activated substrate is the pantetheine aminoacyl group [15]. The crystal structure of the domain that catalyses this reaction has not yet been solved.

Concluding remarks

The folds of CoA-dependent enzymes cover a very diverse set of topologies. The topology seen in the CoA-dependent N-acylating enzymes is unique to this family and thus further increases the diversity of structures among the CoA-dependent enzymes. In addition, sequence or structural motifs that are conserved across all CoA-binding enzymes do not seem to exist. This diversity of fold correlates with the great diversity of functions of these proteins. The wide range of reactions catalysed by the CoA-dependent enzymes is largely due to the increased reactivity of the carbonyl carbon atom and the α-carbon atom of an acyl group when this group is linked to CoA. This increased reactivity is due to the thioester bond and is not dependent on the complex structure of

Figure 3



The structure of AAT (residues 26–161). The orientation is the same as that used in Figure 2. The protein is coloured according to a rainbow colour ramp from the N terminus (N; blue) to the C terminus (C; red). An asterisk (*) marks the small break between the two antiparallel β strands (green). The CoA moiety is shown in stick representation and coloured according to atom type. (The figure was made using the programs MOLSCRIPT [32] and RASTER3D [33].)

CoA. In fact, some enzymes use cysteine–acyl or lipoamide–acyl thioester bonds to perform the same function. Nevertheless, CoA is the predominant acyl group carrier, despite its complex structure. The involvement of CoA in so many enzymatic reactions suggests an early evolutionary origin for this cofactor. Support for this hypothesis lies in the observation that the pantetheine moiety of CoA can be synthesised under prebiotic conditions [16]. The possible early evolutionary origin of CoA may also help to explain the large diversity in the folds of enzymes that catalyse CoA-dependent reactions.

Acknowledgements

It is a pleasure to thank the research groups of Drs Burley, Pauptit and Ramakrishnan for providing access to their coordinates before release by the PDB.

References

1. McGilvery, R.W. (1970). *Biochemistry, a Functional Approach.* pp. 206-262, WB Saunders and Company, Philadelphia, USA.
2. Mahler, H.R. & Cordes, E.H. (1969). In *Biological Chemistry.* pp. 335-339, Harper, Row and Weatherhill, New York, USA.
3. Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem. Rev.* **97**, 2465-2497.
4. Kadonaga, J.T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* **92**, 307-313.
5. Weston, S.A., *et al.*, & Pauptit, R.A. (1998). Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat. Struct. Biol.* **5**, 213-221.
6. Dutnall, R.N., Tafrov, S.T., Sternglas, R. & Ramakrishan, V. (1998). Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell* **94**, 427-438.
7. Neuwald, A.F. & Landsman, D. (1997). GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem. Sci.* **22**, 154-155.
8. Wolf, E., Vassilev, A., Makino, Y., Sali, A., Nakatani, Y. & Burley, S.K. (1998). Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. *Cell* **94**, 439-499.
9. Rossmann, M.G., Liljas, A., Brändén, C.-I. & Banaszak, L.J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes*, *Vol. 11*. (Boyer, P.D., ed), pp. 61-102, Academic Press, New York, USA.
10. Wolodko, W.T., Fraser, M.E., James, M.N.G. & Bridgers, W.A. (1994). The crystal structure of succinyl-CoA synthetase from *Escherichia coli* at 2.5 Å resolution. *J. Biol. Chem.* **269**, 10883-10890.
11. Wierenga, R.K., Terpstra, P. & Hol, W.G.J. (1986). Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* **187**, 101-107.
12. Engel, C.K. & Wierenga, R.K. (1996). The diverse world of CoA-binding proteins. *Curr. Opin. Struct. Biol.* **6**, 790-797.
13. Mathieu, M., *et al.*, & Wierenga, R.K. (1997). The 1.8 Å crystal structure of the dimeric peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: implications for substrate binding and reaction mechanism. *J. Mol. Biol.* **273**, 714-728.
14. Cane, D.E., Walsh, C.T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations and mutations. *Science* **282**, 63-68.
15. Stein, T., *et al.*, & Morris, H. R. (1996). The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *J. Biol. Chem.* **271**, 15428-15435.
16. Keefe, A.D., Newton, G.L. & Miller, S.L. (1995). A possible prebiotic synthesis of pantetheine, a precursor to coenzyme A. *Nature* **373**, 683-685.
17. Remington, S., Wiegand, G. & Huber, R. (1982). Crystallographic refinement and atomic model of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. *J. Mol. Biol.* **158**, 111-152.
18. Mattevi, A., *et al.*, & Hol W.G.J. (1992). Atomic structure of the cubic core of the pyruvate dehydrogenase multienzyme complex. *Science* **255**, 1544-1550.
19. Kim, J.-J.P., Wang, M. & Paschke, R. (1993). Crystal structures of medium-chain acyl-CoA dehydrogenase from pig mitochondria with and without substrate. *Proc. Natl Acad. Sci. USA* **90**, 7523-7527.
20. Mancia, F., *et al.*, & Evans, P.R. (1996). How coenzyme B_{12} radicals are generated: the crystal structure of methylmalonyl-CoA mutase at 2 Å resolution. *Structure* **4**, 339-350.
21. Engel, C.K., Mathieu, M., Zeelen, J.P., Hiltunen, J.K. & Wierenga, R.K. (1996). Crystal structure of enoyl-coenzyme A (CoA) hydratase at 2.5 Å resolution: a spiral fold defines the CoA-binding pocket. *EMBO J.* **15**, 5135-5145.
22. Kragelund, B.B., Andersen, V.K., Madsen, J.C., Knudsen, J. & Poulsen, F.M. (1993). Three-dimensional structure of the complex between acyl-coenzyme A binding protein and palmitoyl-coenzyme A. *J. Mol. Biol.* **230**, 1260-1277.
23. Birktoft, J.J., Holden, H.M., Hamlin, R., Xuong, N.H. & Banaszak, L.J. (1987). Structure of L-3-hydroxyacyl-coenzyme A dehydrogenase: preliminary chain tracing at 2.8 Å resolution. *Proc. Natl Acad. Sci. USA* **84**, 8262-8266.
24. Leslie, A.G.W., Moody, P.C.E. & Shaw, W.V. (1988). Structure of chloramphenicol acetyltransferase at 1.75 Å resolution. *Proc. Natl Acad. Sci. USA* **85**, 4133-4137.
25. Lawrence, C.M., Rodwell, V.W. & Stauffacher, C.V. (1995). Crystal structure of *Pseudomonas mevalonii* HMG-CoA reductase at 3.0 Ångström resolution. *Science* **268**, 1758-1762.
26. Serre, L., Verbree, E.C., Dauter, Z., Stuitje, A.R. & Derewenda, Z.S. (1995). The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5 Å resolution. Crystal structure of a fatty acid synthase component. *J. Biol. Chem.* **270**, 12961-12964.
27. Benning, M.M., *et al.*, & Holden, H.M. (1996). Structure of 4-chlorobenzoyl-coenzyme A dehalogenase determined at 1.8 Å resolution: an enzyme catalyst generated *via* adaptive mutation. *Biochemistry* **35**, 8103-8109.
28. Tiffany, K.A., *et al.*, & Kim, J.J. (1997). Structure of human isovaleryl-CoA dehydrogenase at 2.6 Å resolution: structural basis for substrate specificity. *Biochemistry* **36**, 8455-8464.

29. Lerche, M.H., Kragelund, B.B., Bech, L.M. & Poulsen, F.M. (1997). Barley lipid-transfer protein complexed with palmitoyl-CoA: the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands. *Structure* 5, 291-306.
30. Modis, Y., Filppula, S.A., Novikov, D., Norledge, B., Hiltunen, J.K. & Wierenga, R.K. (1998). The crystal structure of dienoyl-CoA isomerase at 1.5 Å resolution reveals the importance of aspartate and glutamate sidechains for catalysis. *Structure* 6, 957-970.
31. Beaman, T.W., Sugantino, M. & Roderick, S.L. (1998). Structure of the hexapeptide xenobiotic acetyltransferase from *Pseudomonas aeruginosa*. *Biochemistry* 37, 6689-6696.
32. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946-950.
33. Merritt, E.A. & Bacon, D.J. (1997). Raster 3D photorealistic molecular graphics. *Methods Enzymol.* 277, 505-524.